COMPOSITIO MATHEMATICA

J. F. KOKSMA

The theory of asymptotic distribution modulo one

Compositio Mathematica, tome 16 (1964), p. 1-22
<http://www.numdam.org/item?id=CM_1964__16__1_0>

The theory of asymptotic distribution modulo one *

by

J. F. Koksma

§ 1.

The theory of asymptotic distribution (mod 1), having its roots in Kronecker's investigations on the behaviour of the fractional parts of linear forms with integer variables, more or less directly originates from work by Bohl on secular perturbations, Sierpinski on irrational numbers, Borel and F. Bernstein on probability and Hardy-Littlewood on diophantine approximations and Fourier series. Their ideas focussed in Weyl's papers of 1914 and 1916, where by a simple definition, he coined the general notion of *uniform distribution* (mod 1) (equipartition or equidistribution modulo 1) of a sequence of real numbers and as the first gave a necessary and sufficient criterion.

One can 't say that before Weyl in this field there was a lack of ideas which would justify here Goethe's word: "Denn eben wo Begriffe fehlen, da stellt ein Wort zur rechten Zeit sich ein", but nevertheless, it was the introduction of the general notion of uniform distribution (mod 1), be it together with the practicable criterion, which opened the door to a world of new problems. Thus it cannot wonder that after the appearance of Weyl's paper in Mathematische Annalen 77 (1916), his work got an essential impact on the fields where it originated from [1]. If nowadays some one would try to compile a complete bibliography on our subject he would have to check each issue of say the Mathematical Reviews, green covered as well as red covered, as also is shown by the lists of participants and lectures in our symposium and as e.g. one glance in the recent survey by Cigler and Helmberg will make clear [2].

* In 1962 a Nuffic International Summer Session in Science dedicated to "Asymptotic distribution modulo one" was organized at Nijenrode Castle in the Netherlands. The papers presented there will be published collectively in this volume and are all marked "Nijenrode lecture". The above paper aims to give a general introduction to the contemplated field. The papers of this series are quoted here simply by name of author and number of page in this volume.

§ 2.

But this is not all which can be said of Weyl's Annalenpaper of 1916: It is remarkable that several germs (if not to say: *most* germs) of future developments already can be traced there. For, like in many of his publications, with a minimum of calculation Weyl with mastership, indicates several ways to which his ideas give access, obviously without having the need of carrying out further developments himself.

Let me give some examples.

a. Firstly, the *definition*. Take a sequence (u_n):

$$u_1, u_2, \ldots \tag{1}$$

of real numbers and their fractional parts (residues mod 1)

$$\{u_1\}, \{u_2\}, \ldots, \text{ i.e. } u_1-[u_1], u_2-[u_2], \ldots \tag{2}$$

which all are situated in the unit interval

$$E\colon 0 \leqq u < 1 \tag{3}$$

and take a fixed interval

$$\mathscr{I}\colon \alpha \leqq u < \beta;\ \mathscr{I} \subset E. \tag{4}$$

Let $N' = N'(\mathscr{I})$ denote the number of those among the first N numbers (2) which fall into $\mathscr{I}$. *Then if and only if for each fixed* $\mathscr{I} \subset E$:

$$\frac{N'}{N} \to \beta-\alpha \quad \text{as} \quad N \to \infty, \tag{5}$$

the sequence (1) *will be called uniformly distributed* (*mod* 1) (equidistributed mod 1; gleichverteilt mod 1; équirépartie mod 1.)

For N' we may write

$$N' = \sum_{n=1}^{N} \Theta\,(u_n), \tag{6}$$

where $\Theta(u)$ denotes the characteristic function of $\mathscr{I}$ continued periodically along the real axis with period 1. If we put

$$R(N) = R_{\alpha,\beta}(N) = R(N; \mathscr{I}) = N'-(\beta-\alpha)N, \tag{7}$$

(5) asserts that the "*remainder*"

$$R(N) = o(N). \tag{8}$$

Now already from the beginning Weyl shows that if a sequence (1) suffices (5), the formula (5) will hold *uniformly* with respect to the choice of $\mathscr{I}$ in E. Here obviously we have an introduction de facto of the later so called *discrepancy* of a sequence (1), i.e. the number $D(N)$ defined by

$$ND(N) = \sup_{(\mathscr{I})} |N'(\mathscr{I}) - (\beta - \alpha)N| = \sup_{(\mathscr{I})} |R(N; \mathscr{I})| \tag{9}$$

and the proof that for any sequence (1) which is equidistributed (mod 1) automatically holds

$$D(N) \to 0, \text{ if } N \to \infty. \tag{10}$$

It is clear that one has to choose which number one will call discrepancy: one might have given that name also to $ND(N)$ in stead of $D(N)$ and some authors do [3].

The close relation between the discrepancy of a sequence (1) and the way in which the numbers (2) are distributed over E may be illustrated by the following simple formula [4] in which the N numbers $\{u_1\}, \ldots, \{u_N\}$are supposed to be arranged in order of non-decreasing magnitude:

$$\int_0^1 R^2_{0,t}(N)dt = N \sum_{n=1}^{N} \left(\{u_n\} - \frac{n - \frac{1}{2}}{N} \right)^2 + \tfrac{1}{12}; \tag{11}$$

the left hand member is $\leqq (ND(N))^2$ according to (9). We shall return to $D(N)$ later and remark that Hlawka's first lecture (p. 83) is dedicated to this notion, whereas it also plays an essential rôle in his second one on numerical methods (p. 92) and in Erdös's paper (p. 52).

b) Secondly, the *criterion*. It is obvious that reduction modulo 1 of real numbers maps the real axis into the circonference of a circle with radius $1/2\pi$ and replaces the numbers u_n of (1) by the vectors

$$\bar{v}_n = e^{2\pi i u_n}. \tag{12}$$

It is to be expected that the more regularly the numbers u_n are distributed modulo 1 the smaller the absolute value of the sum

$$\sum_{n=1}^{N} \bar{v}_n = \sum_{n=1}^{N} e^{2\pi i u_n} \tag{13}$$

will get for increasing values of $N \to \infty$. (fig. 1).

Weyl's introduction of $e^{2\pi i u}$ by the words: *the proper invariant of the number classes modulo one* implicitly points at notions of greater generality and at abstract analoga, say groups and operators. Several such generalizations have been given later (e.g. Bundgaard, Eckmann) [5]. We return to this topic later. The

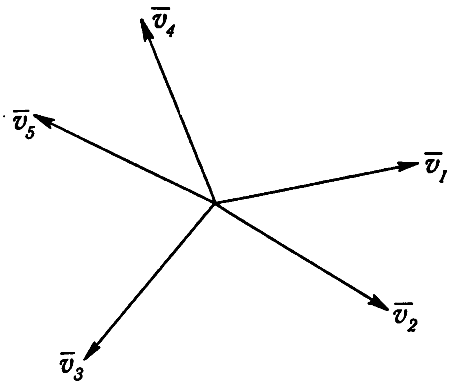

Figure 1

papers of Hartman (p. 66), Helmberg (p. 72, 196), Kemperman (p. 138) will deal with their investigations of this kind.

As to the criterion itself, i.e. the statement that *the relation*

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h u_n} \to 0 \text{ for } N \to \infty \tag{14}$$

for each fixed integer $h \neq 0$ *is sufficient and necessary for the uniform distribution* (*mod* 1) *of the sequence* (1), it can be said that it caused a real "hausse" in the theory of exponential and trigonometrical sums as shows the older work of Hardy-Littlewood, van der Corput, Vinogradoff, Chowla, Walfisz a.o. [6], a topic which is most vital still today and certainly will turn up in several of our symposium lectures.

c) Thirdly, the actual *proof* of the criterion (14) in the one part ("sufficient") gives a link with the theory of Fourier series and in the other part ("necessary") deals with integrals by means of the relation

$$\frac{f(u_1)+\ldots+f(u_N)}{N} \to \int_0^1 f(t)dt, \tag{15}$$

which is derived there for each fixed periodic function $f(x) \in R$ with period 1. Both aspects give way to various investigations.

d) Fourthly, Weyl's first example of a uniformly distributed sequence (1) (mod 1), viz (αn) for each fixed irrational α, with its more dimensional generalizations refines the statements of Kronecker's theorem [7] and will play an important rôle in the analytic theory of numbers (e.g. concerning the Riemann Zeta function) and in Bohr's concept of almost periodic functions. But Weyl moreover derives his inequality, which enables us to jump from (αn) to (αn^2) etc. and is a germ for a lot of later work centering around what van der Corput called the *fundamental inequality* which in its simplest form states *that for a given real function $f(n)$ of the integer n in $a \leqq n < b$ (a and b integers with $b-a = N \geqq 1$) and for each integer ρ ($2 \leqq \rho \leqq N$), one has*

$$(16) \qquad \left|\frac{1}{N}\sum_{n=a}^{b-1} e^{2\pi i f(n)}\right|^2 \leqq \frac{2}{\rho} + \frac{4}{\rho^2}\sum_{\sigma=1}^{\rho-1}\frac{1}{N-\sigma}\left|\sum_{n=a}^{b-\sigma-1} e^{2\pi i (f(n+\sigma)-f(n))}\right|.$$

and from which easily follows that the sequence $f(1), f(2), \ldots$ will be uniformly distributed (mod 1), if the sequence of differences

$$(17) \qquad f(n+h)-f(n) \qquad (n = 1, 2, \ldots)$$

for each fixed integer $h \neq 0$ is uniformly distributed (mod 1), a fact which later by Vinogradoff, van der Corput-Pisot, Cassels also was proved and refined with elementary methods (i.e. without use of the transcendental function $e^{2\pi i u}$) and which recently inspired Hlawka to his investigations of "hereditary properties". Cigler's first paper (p. 29) will be dedicated to this inequality which plays a rôle in several of our papers, a.o. in those of Bertrandias (p. 23) and Kemperman (p. 138).

e) Fifthly, Weyl's application to *mean motion* suggests important opportunities of the theory of uniform distribution for the theory of dynamic systems and other physical domains, especially in its links with probability theory. Here may be pointed at Slater's paper (p. 176).

But his application to mean motion also is important from the standpoint of pure mathematics, as Weyl introduces here the notion of uniform distribution of *functions* $f(t)$ of a real *continuous* variable which is interesting in itself and exactly in our days gives way to various new investigations [8]. Namely I think of L. Kuipers' work.

f) Weyl's theorem that the sequence $(\lambda_n \alpha)$ for each sequence of integers

$$(18) \qquad \lambda_1 < \lambda_2 < \ldots$$

is uniformly distributed (mod 1) *for almost all* α, derived from his criterion (14) by means of the formula

$$(19) \qquad \int_0^1 \left|\sum_{n=1}^{N} e^{2\pi i h\alpha\lambda_n}\right|^2 d\alpha = \sum_{n=1}^{N}\sum_{m=1}^{N}\int_0^1 e^{2\pi i h\alpha(\lambda_n-\lambda_m)}\, d\alpha = N$$

generalizes Borel's notion of *normal numbers* and puts it in the right light: the number N' of those among the first N residues

$\{\alpha 2^n\}$ which are $< \frac{1}{2}$ is precisely the number N' of the zero's among the first N digits in the dyadic expansion of α and thus we get a very special case of an equidistribution problem (mod 1) when we ask for *normal numbers* α, i.e. such numbers in the dyadic expansion of which the number of zero's asymptotically equals the number of ones. Weyl's treatment of this problem opens the way for numerous metrical investigations. His method (which already formerly lead him to the proof of the Riess-Fischer theorem and here to the deduction of (14) from (19)) would later, refined by Menchoff, Plancherel, Rademacher, Erdös, a.o. [9] become the source of many results in the large domain where measure theory, probability theory and ergodic theory meet with number theory, culminating in a steadily increasing variety of cases where the law of the iterated logarithm can be proved to hold [10]. Our speakers de Vroedt (p. 191), Erdös (p. 52), Kemperman (p. 106), Cigler (p. 35), Philipp (p. 161) and Volkmann (p. 186) will touch these subjects. I do not know, whether in 1916 Weyl expected such a spectacular development, exactly in this domain and whether he later would have stuck still to his words, when speaking of "almost-all"-theorems: "Ich glaube dass man den Wert solcher Sätze nicht eben hoch einschätzen darf" [11].

§ 3.

After this tribute to Weyl, let us turn to the definition of uniform distribution as such. The first question which presents itself, is to apply that notion to a given sequence (1). Now it is obvious that if (1) increases slowly, like e.g. the sequences

$$(\log\log n),\ (\log n),\ (\sqrt{n}),\ \ldots \tag{20}$$

the problem will be easier to deal with than if (1) increases more rapidly like the sequences

$$(\theta n^k),\ (e^n),\ (e^{e^n}),\ \ldots; \tag{21}$$

in the first case one can follow the fractional parts so to speak on their paths, but in the latter those small fractional parts mostly are to well hidden behind the large integer parts [12]. Now the theorem (16) and its application based on the behaviour of (17) enables us to move upward, but unfortunately the practical possibilities remain quite restricted, viz mainly to functions $f(n)$ which roughly spoken increase like polynomials

(22) $$f(n) = P(n);\ f(n) = n^{k+\frac{1}{2}}$$

etc. Nevertheless, as van der Corput already proved in the twentieths in his hitherto unpublished part III [13] of "Diophantische Ungleichungen", one can construct entire transcendental functions

(23) $$f(z) = \sum_{h=0}^{\infty} a_h z^h$$

with positive coefficients a_h, such that the sequence

(24) $$f(1), f(2), \ldots$$

is uniformly distributed (mod 1), whereas

(25) $f(x) = x^{\varphi(x)}$ with $\varphi(x) \to \infty$ for positive $x \to \infty$.

But in fact mostly no one knows whether for a given irrational constant θ, like e, π, e^{π}, . . ., the sequence

(26) $$\theta, \theta^2, \theta^3, \ldots$$

fulfills the criterion (14) or not.

On the other hand the theorem that (14) holds for almost all $\theta > 1$ gives a clue and the investigations of Pisot, Salem, Vijajaraghavan and others brought several results on (26). See the joint paper by Pisot-Salem (p. 164). Even for rational $\theta = a/b > 1$ there are various interesting problems here, as e.g. Mahler's work shows [14].

§ 4.

It is obvious that one cannot expect every sequence (1) to fulfil the criterion (14), even if the numbers (2) lay dense in E. Nor even may exist a *distribution function* $\varphi(\gamma)$:

(27) $\varphi(0) = 0,\ \varphi(1) = 1,\ \varphi(\gamma)$ non-decreasing in E

i.e. a function such that for each fixed interval $\mathscr{I} = (0, \gamma) \subset E$

(28) $$\frac{N'}{N} \to \varphi(\gamma), \text{ if } N \to \infty$$

where $N' = N'(\mathscr{I})$ is defined like before (§ 2a with $\alpha = 0$). But in any case we can consider the limits

(29) $$\text{Lim inf}\,\frac{N'}{N} = \varphi(\gamma),\ \text{Lim sup}\,\frac{N'}{N} = \Phi(\gamma) \quad (N \to \infty),$$

φ and Φ being functions of the kind (27) with

(30) $$0 \leqq \varphi \leqq \Phi \leqq 1$$

which are called the *lower* resp. the *upper* distribution function (mod 1) of the sequence (1). If $\varphi \equiv \Phi$ we simply call φ the distribution function, a case thoroughly investigated by Schoenberg [15], who a.o. also proved an analogue of Weyl's criterion: *the sequence* (1) *then and only then has the distribution function* (*mod* 1) $\varphi(\gamma)$, *if*

(31) $$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h u_n} \to \int_0^1 e^{2\pi i h\gamma}\, d\varphi(\gamma), \text{ if } N \to \infty,$$

for each fixed integer $h \neq 0$.

Interesting examples of sequences (1) for which $\varphi < \Phi$ are furnished by slowly increasing functions like loglog x and log x (cf. [12]). Instances of sequences (1) having a continuous distribution function $\varphi(\gamma)$, but not equidistributed (mod 1), are found if one develops certain irrationals θ as a regular continued fraction $\theta = (b_0, b_1, \ldots)$ and considers the sequence of corresponding irrational residues θ_n:

(32) $$\theta = \theta_0 = b_0+\theta_1,\ \theta_1^{-1} = b_1+\theta_2, \ldots,\ \theta_{n-1}^{-1} = b_{n-1}+\theta_n, \ldots$$

This topic will be treated by C. de Vroedt (p. 191). In view of the variety of problems which one meets in cases like these where (2) is by no means equidistributed, for our symposium the name "asymptotic distribution mod 1", was preferred over the name "uniform distribution mod 1."

§ 5.

If one *does not succeed* in proving final results like (5) or (28), one should try to find surrogates; if one *does succeed*, one should not consider (5) or (28) to be really final, but attempt to replace those qualitative formulae by stronger, quantitative, ones.

It therefore may not wonder that there are in existence several variants of the main notion of asymptotic distribution (mod 1), *simplifications*, *complications*, *as* well as *sophistications*. Without trying to give an exhaustive survey, I may list some examples.

a) A first one is to restrict oneself to the simpler question whether for a given sequence (1) the numbers (2) lay everywhere dense in E or not, thus weakening the notion *asymptotic distribution* but still maintaining the words *modulo* 1. Exactly this problem covers the original question of diophantine approximation:

whether a given set of inequalities (mod 1) has infinitely many solutions in integers or only a finite number. In the work of van der Corput and his school around 1930 the investigation of this problem was the main purpose and Weyl's theorem was only a welcome tool which gave more than was asked for [16]. Its applicability is clear: if one knows that for a certain sequence (1) holds (10), then for each integer N and each pair of real numbers α_N and β_N with

$$0 < \beta_N - \alpha_N < 1 \tag{33}$$

the inequality

$$\alpha_N \leqq u_n < \beta_N \pmod 1 \tag{34}$$

has at least one integer solution n in $1 \leqq n \leqq N$, if

$$\beta_N - \alpha_N > D(N), \tag{35}$$

which follows immediately from

$$|N' - (\beta_N - \alpha_N)N| \leqq ND(N), \text{ i.e. } N' > 0, \tag{36}$$

where N' is the number of u_n $(1 \leqq n \leqq N)$ sufficing (34).

Now consider the interesting case

$$\beta_N - \alpha_N \to 0, \text{ if } N \to \infty \tag{37}$$

Then one may try to prove

$$\frac{N'}{(\beta_N - \alpha_N)N} \to 1 \tag{38}$$

in view of (5). A general theorem of van der Corput which he proved already at the end of the twentieths, gives all one may want here, even in the more dimensional case where one considers $s = s(N)$ inequalities

$$\alpha_N^{(\sigma)} \leqq f_\sigma(n) < \beta_N^{(\sigma)} \pmod 1 \qquad (\sigma = 1, 2, \ldots, s) \tag{39}$$

for $s = s(N)$ given sequences $\{f_\sigma(n)\}$ instead of one sequence $\{u_n\}$. The theorem even permits that s turns to infinity with increasing N.

I already said that the method gives more than is asked for, viz (38), that is to say in those cases where it can be applied at all. On the other hand it fails in many cases, where nevertheless the problem itself has a clear sense, as e.g. is shown by van der Corput's theory of rhythmic functions [17]. This restricted applicability of uniform distribution (mod 1) is due a.o. to the fact that

that notion is neither invariant against addition of the considered sequences, nor against other simple operations, e.g. differentiation of the functions $f(x)$ which furnish sequences $\{f(n)\}$ by putting $n = 1, 2, \ldots$.

Also other more direct methods are preferable over Weyl's method and its modifications, if one restricts the problem in the mentioned way. Namely such is the case for the metrical problems which arise when one tries to generalize Khintchine's theorem, that for *almost all real numbers* θ *the inequality*

$$\alpha < \theta n < \alpha + \varphi(n) \pmod 1 \tag{40}$$

has an infinity of integer solutions n, *if* $\sum \varphi(n)$ *diverges and only a finite number of such solutions, if* $\sum \varphi(n)$ *converges.*

Several authors [18] proved that one can replace in (40) the function θn by other functions $f(\theta, n)$. It is remarkable that in contradistinction to "individual" cases most "metrical" problems are easier to deal with when $f(\theta, n)$ for fixed θ grows rapidly with n (lacunary case). De Vries considered with success several non-lacunary cases e.g.

$$\theta P(n), \theta \sqrt[m]{P(n)} \tag{41}$$

where $P(n)$ is a polynomium in n. He moreover dealt with the refined case that the Lebesgue measure is replaced by the Hausdorff dimension [19].

b) A second simplification of totally other kind is to omit the words *modulo one* by assuming from the beginning that the sequence (1) already lies in E, or, as the interval is irrelevant now, in any other interval say (A, B) or even $(-\infty, \infty)$. Then the investigations concern the general idea of distribution functions and although bereft of one of their main charms, viz the arithmetical one, nevertheless may lead to important results, which also in the "modulo 1 case" may find useful applications. I remind you of van der Corput's results in this field [20]; he a.o. proved that any sequence (1) lying dense in E may be ordered in such a way that any pair of distribution functions φ and Φ with (30) may act as lower and upper distribution function of the sequence and similar theorems which once more illustrate that the asymptotical distribution of a sequence is not merely a matter of the set (1) as such, but much more of the way in which it is ordered.

A very hard question which had been posed in 1935 by van der Corput [20], was answered by Mrs van Aardenne-Ehrentfest in 1945. Her answer was refined later by K. Roth [21]. The question is whether a "just", "a democratic" distribution exists; i.e. does a sequence

(42) $$u_1, u_2, \ldots \text{ with } u_n \in E$$

exist, such that for some constant $c_0 > 0$

(43) $$ND(N) \leqq c_0 \qquad (N = 1, 2, \ldots).$$

The answer is "No" and had been conjectured already in view of the fact that the result nearest to (43) that ever had been found is

(44) $$ND(N) \leqq c_1 \log N \qquad (N = 2, 3, \ldots)$$

(e.g. for the special sequences of the form $u_n = \{\theta n\}$, where θ denotes an irrational number such that the partial quotients b_t in its development (32) as a regular continued fraction are bounded.) Roth's proof that (43) can hold for no sequence (42) whatsoever follows from his theorem that for any N-tuple of points $P_n = (u_n, v_n)$ $(n = 1, 2, \ldots, N)$ in the unit square

$$0 \leqq u < 1, \quad 0 \leqq v < 1$$

the relation

(45) $$\int_0^1 \int_0^1 (R(x, y; N))^2 dx dy \geqq c_2 \log N$$

holds, where the absolute constant c_2 does depend neither on N nor on the considered points P_n and where in analogy to (7) $R(x, y; N)$ (the remainder) in this two dimensional case is defined by

(46) $$R(x, y; N) = N' - xyN,$$

N' denoting the number of those among the points $P_1, \ldots, P_N$ which satisfy the simultaneous inequalities

$$0 \leqq u_n < x, \quad 0 \leqq v_n < y \qquad (1 \leqq n \leqq N).$$

Comparing (45) and (11) one sees that there consists a striking difference between the one dimensional case and the two dimensional case. Now in order to study (42), for arbitrary $N \geqq 1$, Roth puts $v_n = n/N$ and remarks that for the N points $(u_n, n/N)$ $(n = 1, 2, \ldots, N)$, the expression (46) differs at most 1 from the expression $R_{0,x}([yN])$ defined by (7) for the numbers $u_1, u_2, \ldots, u_{[yN]}$ from (42). Hence by (45)

(45a) $$|R_{0,x}(M)| \geqq c_2^* \sqrt{\log N}$$

for at most one couple x, M with $0 \leqq x < 1$, $1 \leqq M \leqq N$.

There remains a gap between (44) and (45a). But in any case (45) in a certain sense is best possible as was shown by Davenport [22].

It is to be remarked that dealing with problems of the kind, where $u_n \in E$ (like in (42)) is assumed beforehand, one in Weyl's criterion might replace his sequence of functions

$$(e^{2\pi iht}) \qquad (h \lessgtr 0 \text{ an integer})$$

by any other, say orthonormal and complete sequence

$$(\psi_h(t)) \text{ in } E.$$

In this respect it is to be reminded that various important generalizations of the notion of uniform distribution modulo 1 belong to this category, if from the beginning the space or group concerned as a whole plays the rôle of the unit interval: one might as well omit the words *modulo one* then.

c) Let us return to the real one dimensional case and consider a third modification, which can act as a simplification as well as a complication. I mean the notion of asymptotic distribution (mod 1) with respect to a given sequence of intervals in which the numbers u_n of (1) are defined

$$(47) \qquad a_\sigma \leqq n \leqq b_\sigma \qquad (\sigma = 1, 2, \ldots)$$

in stead of the natural sequence of intervals used originally:

$$(48) \qquad 1 \leqq n \leqq N \qquad (N = 1, 2, \ldots).$$

This conception which from the start systematically was used by van der Corput and his school [23] has some useful features:

α) a fixed sequence (1) automatically being defined in all intervals (47) may have different distribution functions for different sequences (47),

β) one may consider more than 1 sequence (47) at the time, e.g. replace (48) by

$$(49) \qquad m+1 \leqq n \leqq m+N \qquad (m = 1, 2, \ldots).$$

If one asks that $\{u_n\}$ be uniformly distributed in all of the sequences (49) and moreover *uniformly* in m one gets Petersen's definition of a *well distributed sequence* (mod 1), which will be considered in Erdös's paper (p. 52).

γ) One may define u_n in a different way for each individual interval (47). Taking e.g. (48) one may define a function $f_N(n)$ on

(48), which depends on N. Then one defines $\varphi(\gamma)$ and $\Phi(\gamma)$ in the usual way by (29) denoting by $N' = N'(N)$ the number of values n for which $1 \leqq n \leqq N$ and

$$0 \leqq f_N(n) < \gamma \pmod 1.$$

Thus e.g. one may study the asymptotic distribution functions (mod 1) in the intervals (48) of the zero's $f_N(1), f_N(2), \ldots, f_N(N)$ of the N-th eigenfunction $g_N(x)$ of the orthonormal system $\{g_N(x)\}$ belonging to a Sturm-Liouville differential equation with boundary conditions for $0 \leqq x \leqq 1$.

d) A fourth modification, a kind of *weighted* uniform distribution (mod 1) has been introduced by Tsuji [24]. By using the weights

$$\Lambda_1 > 0, \quad \Lambda_2 > 0, \ldots \tag{50}$$

he replaces in the definition (5) the expression

$$\frac{N'}{N} = \frac{\sum_{n=1}^{N} \theta(u_n)}{N} \tag{51}$$

by

$$\frac{\sum_{n=1}^{N} \Lambda_n \Theta_n(u_n)}{\sum_{n=1}^{N} \Lambda_n}. \tag{52}$$

Most known theorems can be carried over then also for this case. This concept can be generalized still further by introducing arbitrary *summability matrices* as has been done by Hlawka and Cigler [25] and about which from the last named a paper will appear in this volume (p. 44).

e) A fifth modification is that of *complete distribution* (mod 1) due to Koroboff [26]: a sequence $(f(n))$ is called completely uniformly distributed (mod 1), if and only if for each set of fixed integers $s \geqq 1$, $a_1, \ldots, a_s \neq 0, \ldots, 0$ the sequence of numbers

$$g(n) = a_1 f(n+1) + \ldots + a_s f(n+s) \qquad (n = 1, 2, \ldots)$$

is uniformly distributed (mod 1). Mark that for $s = h+1$, $a_1 = -1$, $a_s = +1$, $a_\sigma = 0$ $(2 \leqq \sigma \leqq s-1)$ we have the special sequence (17) considered by Weyl and van der Corput.

f) A sixth modification due to LeVeque is the notion of distribution of a sequence (1) modulo a subdivision of the positive real axis [27]. It will be mentioned in Erdös's paper (p. 52).

g) In § 2e I already mentioned the problem of studying the asymptotic distribution (mod. 1) of functions of a continuous variable. Most modifications of the "discrete" theory, also can be applied to the "continuous" case [28].

h) Considering (1), one may study as well the behaviour of the numbers $[u_n]$, instead of $\{u_n\}$. Such is done in Niven's paper (p. 158).

§ 6.

Let us now, returning to the sequence (1) and to the original definition, consider the problem of replacing (8) by preciser results. Then we ask for estimates of

$$|N'-(\beta-\alpha)N| \quad \text{resp. of } D(N), \tag{53}$$

like in § 5b or more general of

$$\left|\sum_{n=1}^{N} f(u_n)-N\int_0^1 f(t)dt\right| \tag{54}$$

for a given (periodic) function $f(t)$ belonging to some integrability class. Now for each $f(t)$ of *bounded variation* the formula

$$\int_0^1 R_{0,t}(N)df(t) = N\int_0^1 f(t)dt - \sum_{n=1}^{N} f(u_n) \tag{55}$$

holds, if suitable interpreted; it learns

$$\left|\sum_{n=1}^{N} f(u_n)-N\int_0^1 f(t)dt\right| \leqq TND(N), \tag{56}$$

where T denotes the total variation of $f(t)$ in $(0, 1)$. Analogously we have

$$\left|\sum_{n=1}^{N} f(u_n)-N\int_0^1 f(t)dt\right| \leqq 3N^2D(N)\omega\left(\frac{1}{N}\right), \tag{57}$$

if f *is continuous* with measure of continuity $\omega(h)$ (i.e. $|f(t+h)-f(t)| \leqq \omega(|h|)$). Applying (56) for $f = e^{2\pi it}$ one finds

$$\left|\sum_{n=1}^{N} e^{2\pi i u_n}\right| \leqq 2\pi ND(N), \tag{58}$$

which gives a lower bound of $ND(N)$ by means of an exponential sum of the Weyl-type.

To find an upper bound for $ND(N)$, one can apply the underlying idea of Weyl's proof of the criterion (14): Approximating the characteristic function $\Theta(u)$ of $\mathcal{J}$ by trigonometric sums, it is possible to express an upper bound for $|R_{\alpha,\beta}(N)|$ in terms of exponential sums of the type occurring in (14).

This idea was worked out in the thirtieths by van der Corput and the author of this paper in a joint proof of a general, more dimensional, theorem [29]: Θ was approximated by the partial sums of the Fourier series of an auxiliar function Θ^* which was constructed in such a way that Θ and Θ^* were identic except near the endpoints α, β (mod 1) and that Θ^* was derivable infinitely often. The theorem has been applied several times.

The idea also was worked out in a different way in 1949 by Erdös and Turán [30]. Instead of introducing an intermediary function Θ^* they used as trigonometric sums the Dunham Jackson means of the discontinuous function $\Theta(u)$ itself and proved a beautiful one dimensional theorem, which is simpler and sharper than ours. The method also proved to be useful to attain a satisfactory more dimensional theorem [31]. In the one dimensional case it runs as follows: *For any sequence* (1), *any* $N \geqq 1$, *any* $M \geqq 1$ *and any* $\mathcal{I} \subset E$ *we have*:

$$(59) \qquad \left|R_{\alpha,\beta}(N)\right| \leqq \frac{150\,N}{M} + \sum_{h=1}^{M} p_h \left|\sum_{n=1}^{N} e^{2\pi i h u_n}\right|,$$

with

$$p_h = \text{Min}\left((\beta-\alpha+\frac{150}{M},\ 1-(\beta-\alpha)+\frac{150}{M},\ \frac{30}{h}\right).$$

§ 7.

In the literature several estimates of expressions (53) or (54) for special classes of sequences can be found, derived by various methods.

Another problem is of methodical significance: to estimate $ND(N)$ by *elementary* methods, i.e. without using transcendental functions like $e^{2\pi i u}$. Here Vinogradoff, van der Corput-Pisot and Cassels deduced inequalities [32] which estimate the discrepancy D of the N numbers $u_1, u_2, \ldots, u_N$ directly by the discrepancy F of the N^2 numbers $u_n - u_m$. Cassels proved [33]

$$D \leqq c_0 F^{\frac{1}{2}}(1+|\log F|).$$

A. Drewes applied elementary methods to various problems in our field [34].

§ 8.

Many investigations concerning $D(N)$ are of a metrical character. I mentioned already the uniform distribution (mod 1) of the special sequence

(60) $(\alpha 2^n)$ for almost all α

and its relation to Borel's notion of normal numbers Borel's statement that *for almost all* α:

(61) $$R(N) = R_{0,\frac{1}{2}}(N) = o(N),$$

which is included in Weyl's theorem of § 2f and has been improved successively [35] by Hausdorff to

(62) $$R(N) = o(N^{\frac{1}{2}+\varepsilon}) \qquad \text{(for each constant } \varepsilon > 0)$$

by Hardy-Littlewood to

(63) $$R(N) = 0(N^{\frac{1}{2}} \log^{\frac{1}{2}} N),\ R(N) = \Omega(N^{\frac{1}{2}}),$$

and by Khintchine firstly to

(64) $$R(N) = 0((N \log \log N)^{\frac{1}{2}})$$

and later to his final estimate [36].

(65) $$\limsup \frac{R(N)}{(N \log \log N)^{\frac{1}{2}}} = \frac{1}{\sqrt{2}}.$$

With these formulae we touch upon probability theory. From the beginning Borel's normal numbers played an essential rôle in his concept of geometric probability and also Khintchine in his proof of (65) used the probabilistic language instead of that of measure theory. His further investigations lead Khintchine to *the law of the iterated logarithm* in probability theory, where it deals with unlimited sequences of Bernoulli trials (each such sequence representing an infinite dyadic fraction, if one denotes success by 1 and failure by 0) with equal probability $\frac{1}{2}$ for success and failure. However, as further work by Khintchine and Kolmogoroff made clear, the same law holds in much more complicated cases concerning random variables, assumed that certain conditions as to their independence are fulfilled; exactly in its full generality it reveals the background of the formulae (61)—(65),which now appear as mere special cases of the main limit theorems of probability theory in their simplest form [37]. Now in view of such magic tools and with regard to the fact that the original proofs of (61)—(65) (in a time where probability theory not yet systematically had been placed on its axiomatic base of measure theory) did ask considerable efforts, the somewhat laming feeling might befall a number theorist that most problems in the field seem to be settled beforehand, that in any case his methods are of rather little use here and thus that no work might be left to him here. But that feeling would be false: Mostly it appears by no means to be easy to prove in concrete cases that the assumptions of probability theory are fulfilled and in fact the kernel of such a problem remains a matter of arithmetical nature. The more so, if indeed the assumptions are not completely satisfied, but nevertheless the law may be proved to hold, as e.g. in the case of *exponential sums*, which was mastered by Erdös and Gál who proved [38]

$$\text{(66)} \qquad \operatorname{Lim\,sup} \frac{\left|\sum_{n=1}^{N} e^{2\pi i \alpha \lambda_n}\right|}{\sqrt{N \operatorname{loglog} N}} = 1 \qquad \text{(for almost all } \alpha\text{)},$$

where $\lambda_1 < \lambda_2 < \ldots$ is an arbitrarily fixed *lacunary sequence of positive numbers.*

The significance of results of this kind in the light of Weyl's criterion (14) and of the theorem (59) is clear.

§ 9.

A similar situation exists with respect to the *ergodic theory* and the problem of approximating an integral

$$\text{(67)} \qquad \int_0^1 f(t)dt$$

by sums

$$\text{(68)} \qquad \frac{1}{N}\left(f(u_1) + \ldots + f(u_N)\right)$$

where $u_1, u_2, \ldots$ is supposed to be equidistributed (mod 1) and $f(t)$ to be periodic with period 1.

We know that if $f(t) \in R$, the relation (15) surely holds. But as the numbers $\{u_n\}$ in E only form a nullset, one cannot expect that (15) also holds for each $f(t) \in L$. It however becomes a different question, if $u_n = u_n(\theta)$ also depends on a parameter θ, which e.g. on the real axis ranges from A to B and if one asks whether (15) holds almost everywhere (a.e.) in the segment $A \leqq \theta \leqq B$.

Thus Raikoff proved that if $f \in L$

$$\text{(69)} \qquad \frac{1}{N}\sum_{1}^{N} f(\theta a^n) \to \int_0^1 f(t)dt \qquad \text{a.e.,}$$

if $a \geqq 2$ is a fixed integer. But as M. Riess pointed out this result is an instance of the so called individual ergodic theorem [39]. The individual ergodic theorem and its number theoretical applications form the subject matter of one of Cigler's papers (p. **35**). The same individual ergodic theorem is the source of two results of Khintchine [40], resp. Ryll-Nardzewski [41]. Khintchine proved for $f \in L$

$$\frac{1}{N}\sum_{n=1}^{N} f(\alpha n+\theta) \to \int_0^1 f(t)dt \qquad \text{a.e.} \tag{70}$$

if α is a fixed irrational. Ryll-Nardzewski proved

$$\frac{1}{N}\sum_{n=1}^{N} f(\theta_n) \to \frac{1}{\log 2}\int_0^1 f(t)\frac{dt}{1+t} \qquad \text{a.e.} \tag{71}$$

where for each real θ the sequence $\theta_n = \theta_n(\theta)$ is defined by (32).

Now it is remarkable that for other, rather simple, functions $u_n(\theta)$ instead of those appearing in (69), (70), (71) one did not succeed till now in proving (15) for $f(t) \in L$. Thus the case $u_n(\theta) = \theta n$ which already long ago was investigated by Khintchine [42] who even in the simplified case that $f(x)$ is the characteristic function of a measurable set $S \subset E$ only could prove (15) under some additional conditions as to the nature of S. Also one might conjecture that (70) would remain true, if one would replace (αn) there by an arbitrary sequence (u_n) which is equidistributed (mod 1) sufficiently regularly, i.e. for which the discrepancy $D(N)$ would turn to 0 very fast with increasing N. But also here one meets with serious difficulties, which appear to be essential in the light of some important counter examples, due to Erdös [43]:

1) Take $\{u_n\}$ uniformly distributed (mod 1) with a discrepancy as small as possible. Then there exists always a function $f(t) \in L$ even bounded (and thus $\in L^p$; $p \geqq 1$) such that

$$\int_0^1 f(t)dt = 0; \ \text{Lim sup}\, \frac{1}{N}\sum_{n=1}^{N} f(u_n+\theta) > \tfrac{1}{2} \ \textit{for all } \theta. \tag{72}$$

2) There exists a sequence of integers $\lambda_1 < \lambda_2 < \lambda_3 < \ldots$, even lacunary, together with a function $f(x) \in L^p$ ($p \geqq 1$ arbitrary), such that almost everywhere

$$\int_0^1 f(t)dt = 0; \ \text{Lim sup}\, \frac{1}{N}\sum_{n=1}^{N} f(\theta\lambda_n) = \infty. \tag{73}$$

Although in the light of these counter examples several generalizations of (69) and (70) certainly are impossible, they do not cover the rather simple case

$$u_n(\theta) = \theta n,$$

I do not know a final result as to the question for which $f(t)$ the relation

$$\frac{1}{N}\sum_{n=1}^{N} f(\theta n) \to \int_0^1 f(t)dt \qquad \text{a.e.}$$

is true [44], and also in the cases $u_n(\theta) = \theta\lambda_n$, $u_n(\theta) = u_n + \theta$ there remain several unsolved problems. Cf. also Erdös's paper [3].

I remark that in probabilistic language (69), (70) and (71) are instances of the *law of large numbers*. One might try to prove stronger results by estimating the error

$$(74) \qquad \left|\frac{1}{N}\sum_{n=1}^{N} f(u_n(\theta)) - \int_0^1 f(t)dt\right|$$

and casually hope for an estimate according to the law of the iterated logarithm.

§ 10.

As I already mentioned, several papers will be dedicated to the more abstract generalizations of asymptotic distribution modulo one.

A first kind of generalization is to replace the reduction (mod 1) by other operators $T(x)$ operating on the real numbers or on the complex numbers and which preferably form a measure preserving group. One may use this concept to develop a theory of great generality, but also one can pose concrete problems here: Take the upper half plane H of the complex z-plane and its division by a modular group with main domain D_0. Reduce the numbers of a given sequence $z_1, z_2, \ldots \in H$ by taking the representatives $\{z_1\}, \{z_2\}, \ldots \in D_0$ and study their distribution there.

An other kind of generalization is to replace the real numbers themselves by an abstract space or group in which a measure or integral can be defined. Here also various concrete problems may be posed like for the field of P-adic numbers a, where each number can be developed in a Laurent-series

$$(75) \quad a = a_{-s}P^{-s} + \ldots + a_{-1}P^{-1} + a_0 + a_1P + a_2P^2 + \ldots$$

and where e.g. as reduction may be chosen: replacing of a by

$$(76) \qquad \{a\} = a_0 + a_1P + a_2P^2 + \ldots$$

But of course all such concepts may be varied and generalized in numerous ways. In my eyes preferably the arithmetical underground of Weyl's conception should be preserved.

REFERENCES and NOTES

[1] For references till 1936 cf. my *Diophantische Approximationen* (Berlin 1936, Ergebnisse der Mathematik IV, 4), in the following denoted by D. A.

[2] J. Cigler und G. Helmberg, Neuere Entwicklungen der Theorie der Gleichverteilung. Jahresbericht der D.M.V. 64, 1–50 (1961).

A large part of: J. W. S. Cassels, An introduction to diophantine approximation (Cambridge Univ. Tract 45, 1957) also is dedicated to our subject.

[3] E.g. P. Erdös in his contribution to this symposium: Problems and results on diophantine approximations, (this volume p. 52).

[4] For this and similar formulae cf. my notes:

Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo 1. Mathematica (Zutphen) IIB, 7–11 (1942/43).

Eenige integralen in de theorie der gelijkmatige verdeeling modulo 1. Mathematica (Zutphen) 11B, 49–52 (1942/43).

[5] S. Bundgaard, Ueber de Werteverteilung der Charaktere abelscher Gruppen, Math.-fys. Medd. Danske Vid. Selsk. 14, No. 4, 1–29 (1936).

The author bases his work on von Neumann's notion of the mean value of an almost periodic function in a group (Transactions Amer. Math. Soc. 36, 445–492 (1934)).

B. Eckmann, Über monothetische Gruppen. Comment. math. Helvet. 16, 249–263 (1943/44).

[6] Cf. D. A. Ch's VIII, IX.

[7] Cf. D. A. Ch's VII, VIII, IX, also for the remainder of § 2d.

[8] E.g. by L. Kuipers and B. Meulenbeld. For references cf. Cigler-Helmberg quoted in [2].

[9] For references cf. I. S. Gál-J. F. Koksma, Sur l'ordre de grandeur des fonctions sommables, Proc. Kon. Ned. Akad. Wet. 53, 638–653 (1950) = Indagationes Mathematicae 12, 192–207 (1950).

[10] E.g. cf. P. Erdös-I. S. Gál, On the law of the iterated logarithm, Proc. Kon. Ned. Akad. Wet. 58, 65–84 (1955) = Indagationes Mathematicae 17, 65–84 (1955).

[11] In his paper: Über die Gleichverteilung von Zahlen modulo Eins, Math. Ann. 77, 313–352 (1916) p. 345.

[12] Cf. my paper: Asymptotische verdeling van reële getallen modulo 1 I, II, III, Mathematica (Leiden) 1 (1933), 245–248, 2 (1933), 1–6, 3 (1933), 107–114 and D. A. Ch. VIII.

[13] Part I (Zur Gleichverteilung modulo Eins) and Part II (Rhythmische Systeme, A und B) appeared in the Acta Math: J. G. van der Corput, Diophantische Ungleichungen, Acta Math. 56, 373–456 (1931), resp. 59, 209–328 (1932).

[14] K. MAHLER, On the fractional parts of the powers of a rational number, I, Acta Arithm, 3 (1938), 89—93, II, Mathematika (London) 4 (1957), 122—124. For further references concerning (26) etc. cf. the paper of PISOT-SALEM in this volume (p. 164).

[15] I. SCHOENBERG, Ueber die asymptotische Verteilung reeller Zahlen mod. 1. Math. Z. 28, 171—199 (1928).

[16] Cf [13] and D. A. Ch's VIII, IX.

[17] Cf [13].

[18] R. J. DUFFIN and A. C. SCHAEFFER, Khintchine's problems in metric Diophantine approximation. Duke Math. J. 8, 243—255 (1941).
J. F. KOKSMA, Niet-lineaire simultane approximaties. Handel. Ned. Nat. Congres, 95—96 (1941).
ibid. Sur la theorie métrique des approximations diophantiques, Proc. Ned. Akad. Wet. 48, 249—265 (1945). Indagationes Mathematicae 7, 54—70 (1945), where also further references are given.
J. W. S. CASSELS, Some metrical theorems in diophantine approximation. I Proc. Cambr. Phil. Soc. 46, 209—218 (1949). II J. London Math. Soc. 25, 180—184 (1950).

[19] D. DE VRIES, Metrische onderzoekingen van Diophantische benaderingsproblemen in het niet-lacunaire geval. (Diss. Amsterdam, V.U.), 1955.

[20] J. G. VAN DER CORPUT, Verteilungsfunktionen. Proc. Kon. Ned. Akad. Wet. 38, 813—821; 1058—1066 (1938); 39, 10—19; 19—26; 149—153; 339—344; 489—494; 579—590 (1939).

[21] For references cf. K. ROTH, On irregularities of distribution. Mathematika (London) 1, 73—79 (1954).

[22] H. DAVENPORT, Note on irregularities of distribution. Mathematika (London), 3, 131—135 (1956).

[23] Cf. [13] and D. A. Ch's VIII and IX.

[24] M. TSUJI, On the uniform distribution of numbers (mod. 1). J. Math. Soc. Japan 4, 313—322 (1952).

[25] For referènces cf. Dr. CIGLER's third paper in this vol. (p. 44).

[26] N. M. KOROBOFF, Einige Probleme der Verteilung von Bruchteilen. Uspechi mat. Nauk 4, 189—190 (1949).

[27] W. LEVEQUE, On uniform distribution modulo a subdivision. Pacific J. of Math. 3, 757—771 (1953).

[28] In this respect I mention a result by C. RYLL NARDZEWSKI, Sur les suites et les fonctions également réparties. Studia math. 12, 143—144 (1951) which in certain cases gives a link between both theories.

[29] It is the theorem which in its one dimensional case is quoted as Satz 4 in D.A. p. 101 and which itself is related to the old theorem of VAN DER CORPUT, which is meant in § 5a after (38) in this paper. For further references cf also [31]. Several applications a.o. are given by A. DREWES, Diophantische Benaderingsproblemen. (Diss. Amsterdam V.U.), 1945.

[30] P. ERDÖS and A. TURÁN, On a problem in the theory of uniform distribution I, II. Proc. Kon. Ned. Akad. Wet. (ser. A.) 51, 370—378; 406—413 (1948), = Indagationes Mathematicae 10, 370—378; 406—413 (1948).

[31] J. F. KOKSMA, Some theorems on Diophantine inequalities. Scriptum 5 of the Mathematical Centre, Amsterdam (1950).

[32] Cf. D. A. Ch. VIII, IX.

[33] J. W. S. CASSELS, A new inequality with application to the theory of diophantine approximation. Math. Ann. 126, 108—118 (1953).

[34] Cf. his thesis quoted in [29].

[35] Cf. D. A. IX, § 6, p. 116.
Similar problems for generalized dyadic fractions have been treated by C. SANDERS, Verdelingsproblemen bij gegeneraliseerde duale breuken. (Diss. Amsterdam V.U.), 1950.

[36] A. KHINTCHINE, Asymptotische Gesetze der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik II, 4, (1933).

[37] Cf. [36] and e.g. W. FELLER, An introduction to probability theory and its applications I, sec. ed. New York-London (1960).

[38] P. ERDÖS and I. S. GÁL, On the law of the iterated logarithm I, II. Proc. Kon. Ned. Akad. Wet. (ser. A), 58, 64—84 (1955), Indagationes Mathematicae 17, 64—84 (1955).

[39] For ref. cf. e.g. my paper An arithmetical property of some sommable functions. Proc. Kon. Ned. Akad. Wet. (ser. A) 53, 960—972 (1950) = Indagationes Mathematicae 12, 354—367 (1950).

[40] A. KHINTCHINE, Eine arithmetische Eigenschaft der summierbaren Funktionen. Recueil Math., Moscou 41, 11—13 (1934).

[41] C. RYLL-NARDZEWSKI, On the ergodic theorems, I, II. Studia Mathematica XII, 65—79 (1951).

[42] Cf. [39].

[43] A proof of the first counter example in J. F. KOKSMA-R. SALEM, Uniform distribution and Lebesgue integration. Acta Scient. Math. Szeged 12, 87—96 (1950).
A proof of the second counter example in P. ERDÖS, On the strong law of large numbers. Transactions Amer. Math. Soc. 67, 51—56 (1950).

[44] Cf. my paper: Sur les suites $(\lambda_n x)$ et les fonctions $g(t) \in L^{(2)}$. J. de Math. p. appl. 35, 289—296 (1956).

(Oblatum 29-5-63).